

Re-run

T. Saidha



RAW SEQUENCE LISTING

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:06:59

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw

1 <110> APPLICANT: Schroeder, Hartwig
 2 <120> TITLE OF INVENTION: Process for preparing biotin
 3 <130> FILE REFERENCE: 0000048792/UP
 4 <140> CURRENT APPLICATION NUMBER: US/09/622,419
 5 <141> CURRENT FILING DATE: 2003-03-13
 6 <150> PRIOR APPLICATION NUMBER: PCT/EP 99/01052
 7 <151> PRIOR FILING DATE: 1999-02-17
 8 <160> NUMBER OF SEQ ID NOS: 17
 9 <170> SOFTWARE: PatentIn Release 1.0, version 1.25 (EPO)/WordPerfect
 11 <210> SEQ ID NO: 1
 12 <211> LENGTH: 1155
 13 <212> TYPE: DNA
 14 <213> ORGANISM: Escherichia coli
 15 <220> FEATURE:
 16 <221> NAME/KEY: CDS
 17 <222> LOCATION: 1..1155
 18 <400> SEQUENCE: 1

19	atg gca aaa cac ctt ttt acg tcc gag tcc gtc tct gaa ggg cat cct	48
20	Met Ala Lys His Leu Phe Thr Ser Glu Ser Val Ser Glu Gly His Pro	
21	1 5 10 15	
22	gac aaa att gct gac caa att tct gat gcc gtt tta gac gcg atc ctc	96
23	Asp Lys Ile Ala Asp Gln Ile Ser Asp Ala Val Leu Asp Ala Ile Leu	
24	20 25 30	
25	gaa cag gat ccg aaa gca cgc gtt gct tgc gaa acc tac gta aaa acc	144
26	Glu Gln Asp Pro Lys Ala Arg Val Ala Cys Glu Thr Tyr Val Lys Thr	
27	35 40 45	
28	ggc atg gtt tta gtt ggc ggc gaa atc acc acc agc gcc tgg gta gac	192
29	Gly Met Val Leu Val Gly Gly Glu Ile Thr Thr Ser Ala Trp Val Asp	
30	50 55 60	
31	atc gaa gag atc acc cgt aac acc gtt cgc gaa att ggc tat gtg cat	240
32	Ile Glu Glu Ile Thr Arg Asn Thr Val Arg Glu Ile Gly Tyr Val His	
33	65 70 75 80	
34	tcc gac atg ggc ttt gac gct aac tcc tgt gcg gtt ctg agc gct atc	288
35	Ser Asp Met Gly Phe Asp Ala Asn Ser Cys Ala Val Leu Ser Ala Ile	
36	85 90 95	
37	ggc aaa cag tct cct gac atc aac cag ggc gtt gac cgt gcc gat ccg	336
38	Gly Lys Gln Ser Pro Asp Ile Asn Gln Gly Val Asp Arg Ala Asp Pro	
39	100 105 110	
40	ctg gaa cag ggc gcg ggt gac cag ggt ctg atg ttt ggc tac gca act	384
41	Leu Glu Gln Gly Ala Gly Asp Gln Gly Leu Met Phe Gly Tyr Ala Thr	
42	115 120 125	
43	aat gaa acc gac gtg ctg atg cca gca cct atc acc tat gca cac cgt	432
44	Asn Glu Thr Asp Val Leu Met Pro Ala Pro Ile Thr Tyr Ala His Arg	

ENTERED

RAW SEQUENCE LISTING

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:06:59

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw

```

45          130          135          140
46      ctg gta cag cgt cag gct gaa gtg cgt aaa aac ggc act ctg ccg tgg      480
47      Leu Val Gln Arg Gln Ala Glu Val Arg Lys Asn Gly Thr Leu Pro Trp
48      145          150          155          160
49      ctg cgc ccg gac gcg aaa agc cag gtg act ttt cag tat gac gac ggc      528
50      Leu Arg Pro Asp Ala Lys Ser Gln Val Thr Phe Gln Tyr Asp Asp Gly
51          165          170          175
52      aaa atc gtt ggt atc gat gct gtc gtg ctt tcc act cag cac tct gaa      576
53      Lys Ile Val Gly Ile Asp Ala Val Val Leu Ser Thr Gln His Ser Glu
54          180          185          190
55      gag atc gac cag aaa tcg ctg caa gaa gcg gta atg gaa gag atc atc      624
56      Glu Ile Asp Gln Lys Ser Leu Gln Glu Ala Val Met Glu Glu Ile Ile
57          195          200          205
58      aag cca att ctg ccc gct gaa tgg ctg act tct gcc acc aaa ttc ttc      672
59      Lys Pro Ile Leu Pro Ala Glu Trp Leu Thr Ser Ala Thr Lys Phe Phe
60      210          215          220
61      atc aac ccg acc ggt cgt ttc gtt atc ggt ggc cca atg ggt gac tgc      720
62      Ile Asn Pro Thr Gly Arg Phe Val Ile Gly Gly Pro Met Gly Asp Cys
63      225          230          235          240
64      ggt ctg act ggt cgt aaa att atc gtt gat acc tac ggc ggc atg gcg      768
65      Gly Leu Thr Gly Arg Lys Ile Ile Val Asp Thr Tyr Gly Gly Met Ala
66          245          250          255
67      cgt cac ggt ggc ggt gca ttc tct ggt aaa gat cca tca aaa gtg gac      816
68      Arg His Gly Gly Gly Ala Phe Ser Gly Lys Asp Pro Ser Lys Val Asp
69          260          265          270
70      cgt tcc gca gcc tac gca gca cgt tat gtc gcg aaa aac atc gtt gct      864
71      Arg Ser Ala Ala Tyr Ala Ala Arg Tyr Val Ala Lys Asn Ile Val Ala
72          275          280          285
73      gct ggc ctg gcc gat cgt tgt gaa att cag gtt tcc tac gca atc ggc      912
74      Ala Gly Leu Ala Asp Arg Cys Glu Ile Gln Val Ser Tyr Ala Ile Gly
75          290          295          300
76      gtg gct gaa ccg acc tcc atc atg gta gaa act ttc ggt act gag aaa      960
77      Val Ala Glu Pro Thr Ser Ile Met Val Glu Thr Phe Gly Thr Glu Lys
78      305          310          315          320
79      gtg cct tct gaa caa ctg acc ctg ctg gta cgt gag ttc ttc gac ctg      1008
80      Val Pro Ser Glu Gln Leu Thr Leu Leu Val Arg Glu Phe Phe Asp Leu
81          325          330          335
82      cgc cca tac ggt ctg att cag atg ctg gat ctg ctg cac ccg atc tac      1056
83      Arg Pro Tyr Gly Leu Ile Gln Met Leu Asp Leu Leu His Pro Ile Tyr
84          340          345          350
85      aaa gaa acc gca qca tac ggt cac ttt ggt cgt gaa cat ttc ccg tgg      1104
86      Lys Glu Thr Ala Ala Tyr Gly His Phe Gly Arg Glu His Phe Pro Trp
87          355          360          365
88      gaa aaa acc gac aaa gcg cag ctg ctg cgc gat gct gcc ggt ctg aag      1152
89      Glu Lys Thr Asp Lys Ala Gln Leu Leu Arg Asp Ala Ala Gly Leu Lys
90      370          375          380
91      taa      1155
93 <210> SEQ ID NO: 2
94 <211> LENGTH: 384

```

RAW SEQUENCE LISTING

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:06:59

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw

95 <212> TYPE: PRT

96 <213> ORGANISM: Escherichia coli

97 <400> SEQUENCE: 2

```

98 Met Ala Lys His Leu Phe Thr Ser Glu Ser Val Ser Glu Gly His Pro
99      1          5          10          15
100 Asp Lys Ile Ala Asp Gln Ile Ser Asp Ala Val Leu Asp Ala Ile Leu
101      20          25          30
102 Glu Gln Asp Pro Lys Ala Arg Val Ala Cys Glu Thr Tyr Val Lys Thr
103      35          40          45
104 Gly Met Val Leu Val Gly Gly Glu Ile Thr Thr Ser Ala Trp Val Asp
105      50          55          60
106 Ile Glu Glu Ile Thr Arg Asn Thr Val Arg Glu Ile Gly Tyr Val His
107      65          70          75          80
108 Ser Asp Met Gly Phe Asp Ala Asn Ser Cys Ala Val Leu Ser Ala Ile
109      85          90          95
110 Gly Lys Gln Ser Pro Asp Ile Asn Gln Gly Val Asp Arg Ala Asp Pro
111      100         105         110
112 Leu Glu Gln Gly Ala Gly Asp Gln Gly Leu Met Phe Gly Tyr Ala Thr
113      115         120         125
114 Asn Glu Thr Asp Val Leu Met Pro Ala Pro Ile Thr Tyr Ala His Arg
115      130         135         140
116 Leu Val Gln Arg Gln Ala Glu Val Arg Lys Asn Gly Thr Leu Pro Trp
117      145         150         155         160
118 Leu Arg Pro Asp Ala Lys Ser Gln Val Thr Phe Gln Tyr Asp Asp Gly
119      165         170         175
120 Lys Ile Val Gly Ile Asp Ala Val Val Leu Ser Thr Gln His Ser Glu
121      180         185         190
122 Glu Ile Asp Gln Lys Ser Leu Gln Glu Ala Val Met Glu Glu Ile Ile
123      195         200         205
124 Lys Pro Ile Leu Pro Ala Glu Trp Leu Thr Ser Ala Thr Lys Phe Phe
125      210         215         220
126 Ile Asn Pro Thr Gly Arg Phe Val Ile Gly Gly Pro Met Gly Asp Cys
127      225         230         235         240
128 Gly Leu Thr Gly Arg Lys Ile Ile Val Asp Thr Tyr Gly Gly Met Ala
129      245         250         255
130 Arg His Gly Gly Gly Ala Phe Ser Gly Lys Asp Pro Ser Lys Val Asp
131      260         265         270
132 Arg Ser Ala Ala Tyr Ala Ala Arg Tyr Val Ala Lys Asn Ile Val Ala
133      275         280         285
134 Ala Gly Leu Ala Asp Arg Cys Glu Ile Gln Val Ser Tyr Ala Ile Gly
135      290         295         300
136 Val Ala Glu Pro Thr Ser Ile Met Val Glu Thr Phe Gly Thr Glu Lys
137      305         310         315         320
138 Val Pro Ser Glu Gln Leu Thr Leu Leu Val Arg Glu Phe Phe Asp Leu
139      325         330         335
140 Arg Pro Tyr Gly Leu Ile Gln Met Leu Asp Leu Leu His Pro Ile Tyr
141      340         345         350
142 Lys Glu Thr Ala Ala Tyr Gly His Phe Gly Arg Glu His Phe Pro Trp
143      355         360         365

```

RAW SEQUENCE LISTING

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:06:59

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw

```

144      Glu Lys Thr Asp Lys Ala Gln Leu Leu Arg Asp Ala Ala Gly Leu Lys
145          370                      375                      380
147 <210> SEQ ID NO: 3
148 <211> LENGTH: 1206
149 <212> TYPE: DNA
150 <213> ORGANISM: Escherichia coli
151 <220> FEATURE:
152 <221> NAME/KEY: CDS
153 <222> LOCATION: 1..1206
154 <400> SEQUENCE: 3
155      atg aac gtt ttt aat ccc gcg cag ttt cgc gcc cag ttt ccc gca cta      48
156      Met Asn Val Phe Asn Pro Ala Gln Phe Arg Ala Gln Phe Pro Ala Leu
157          1                      5                      10                      15
158      cag gat gcg ggc gtc tat ctc gac agc gcc gcg acc gcg ctt aaa cct      96
159      Gln Asp Ala Gly Val Tyr Leu Asp Ser Ala Ala Thr Ala Leu Lys Pro
160          20                      25                      30
161      gaa gcc gtg gtt gaa gcc acc caa cag ttt tac agt ctg agc gcc gga      144
162      Glu Ala Val Val Glu Ala Thr Gln Gln Phe Tyr Ser Leu Ser Ala Gly
163          35                      40                      45
164      aac gtc cat cgc agc cag ttt gcc gaa gcc caa cgc ctg acc gcg cgt      192
165      Asn Val His Arg Ser Gln Phe Ala Glu Ala Gln Arg Leu Thr Ala Arg
166          50                      55                      60
167      tat gaa gct gca cga gag aaa gtg gcg caa tta ctg aat gca ccg gat      240
168      Tyr Glu Ala Ala Arg Glu Lys Val Ala Gln Leu Leu Asn Ala Pro Asp
169          65                      70                      75                      80
170      gat aaa act atc gtc tgg acg cgc gcc acc act gaa tcc atc aac atg      288
171      Asp Lys Thr Ile Val Trp Thr Arg Gly Thr Thr Glu Ser Ile Asn Met
172          85                      90                      95
173      gtg gca caa tgc tat gcg cgt ccg cgt ctg caa ccg ggc gat gag att      336
174      Val Ala Gln Cys Tyr Ala Arg Pro Arg Leu Gln Pro Gly Asp Glu Ile
175          100                     105                     110
176      att gtc agc gtg gca gaa cac cac gcc aac ctc gtc ccc tgg ctg atg      384
177      Ile Val Ser Val Ala Glu His His Ala Asn Leu Val Pro Trp Leu Met
178          115                     120                     125
179      gtc gcc caa caa act gga gcc aaa gtg gtg aaa ttg ccg ctt aat gcg      432
180      Val Ala Gln Gln Thr Gly Ala Lys Val Val Lys Leu Pro Leu Asn Ala
181          130                     135                     140
182      cag cga ctg ccg gat gtc gat ttg ttg cca gaa ctg att act ccc cgt      480
183      Gln Arg Leu Pro Asp Val Asp Leu Leu Pro Glu Leu Ile Thr Pro Arg
184          145                     150                     155                     160
185      agt cgg att ctg gcg ttg ggt cag atg tcg aac gtt act ggc ggt tgc      528
186      Ser Arg Ile Leu Ala Leu Gly Gln Met Ser Asn Val Thr Gly Gly Cys
187          165                     170                     175
188      ccg gat ctg gcg cga gcg att acc ttt gct cat tca gcc ggg atg gtg      576
189      Pro Asp Leu Ala Arg Ala Ile Thr Phe Ala His Ser Ala Gly Met Val
190          180                     185                     190
191      gtg atg gtt gat ggt gct cag ggg gca gtg cat ttc ccc gcg gat gtt      624
192      Val Met Val Asp Gly Ala Gln Gly Ala Val His Phe Pro Ala Asp Val
193          195                     200                     205

```

RAW SEQUENCE LISTING

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:06:59

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw

```

194   cag caa ctg gat att gat ttc tat gct ttt tca ggt cac aaa ctg tat      672
195   Gln Gln Leu Asp Ile Asp Phe Tyr Ala Phe Ser Gly His Lys Leu Tyr
196       210                               215                               220
197   ggc ccg aca ggt atc ggc gtg ctg tat ggt aaa tca gaa ctg ctg gag      720
198   Gly Pro Thr Gly Ile Gly Val Leu Tyr Gly Lys Ser Glu Leu Leu Glu
199       225                               230                               235                               240
200   gcg atg tcg ccc tgg ctg ggc ggc ggc aaa atg gtt cac gaa gtg agt      768
201   Ala Met Ser Pro Trp Leu Gly Gly Gly Lys Met Val His Glu Val Ser
202       245                               250                               255
203   ttt gac ggc ttc acg act caa tct gcg ccg tgg aaa ctg gaa gct gga      816
204   Phe Asp Gly Phe Thr Thr Gln Ser Ala Pro Trp Lys Leu Glu Ala Gly
205       260                               265                               270
206   acg cca aat gtc gct ggt gtc ata gga tta agc gcg gcg ctg gaa tgg      864
207   Thr Pro Asn Val Ala Gly Val Ile Gly Leu Ser Ala Ala Leu Glu Trp
208       275                               280                               285
209   ctg gca gat tac gat atc aac cag gcc gaa agc tgg agc cgt agc tta      912
210   Leu Ala Asp Tyr Asp Ile Asn Gln Ala Glu Ser Trp Ser Arg Ser Leu
211       290                               295                               300
212   gca acg ctg gcg gaa gat gcg ctg gcg aaa cgt ccc ggc ttt cgt tca      960
213   Ala Thr Leu Ala Glu Asp Ala Leu Ala Lys Arg Pro Gly Phe Arg Ser
214       305                               310                               315                               320
215   ttc cgc tgc cag gat tcc agc ctg ctg gcc ttt gat ttt gct ggc gtt      1008
216   Phe Arg Cys Gln Asp Ser Ser Leu Leu Ala Phe Asp Phe Ala Gly Val
217       325                               330                               335
218   cat cat agc gat atg gtg acg ctg ctg gcg gag tac ggt att gcc ctg      1056
219   His His Ser Asp Met Val Thr Leu Leu Ala Glu Tyr Gly Ile Ala Leu
220       340                               345                               350
221   cgg gcc ggg cag cat tgc gct cag ccg cta ctg gca gaa tta ggc gta      1104
222   Arg Ala Gly Gln His Cys Ala Gln Pro Leu Leu Ala Glu Leu Gly Val
223       355                               360                               365
224   acc ggc aca ctg cgc gcc tct ttt gcg cca tat aat aca aag agt gat      1152
225   Thr Gly Thr Leu Arg Ala Ser Phe Ala Pro Tyr Asn Thr Lys Ser Asp
226       370                               375                               380
227   gtg gat gcg ctg gtg aat gcc gtt gac cgc gcg ctg gaa tta ttg gtg      1200
228   Val Asp Ala Leu Val Asn Ala Val Asp Arg Ala Leu Glu Leu Leu Val
229       385                               390                               395                               400
230   gat taa                                                                    1206
231   Asp
233 <210> SEQ ID NO: 4
234 <211> LENGTH: 401
235 <212> TYPE: PRT
236 <213> ORGANISM: Escherichia coli
237 <400> SEQUENCE: 4
238   Met Asn Val Phe Asn Pro Ala Gln Phe Arg Ala Gln Phe Pro Ala Leu
239       1                               5                               10                               15
240   Gln Asp Ala Gly Val Tyr Leu Asp Ser Ala Ala Thr Ala Leu Lys Pro
241       20                               25                               30
242   Glu Ala Val Val Glu Ala Thr Gln Gln Phe Tyr Ser Leu Ser Ala Gly
243       35                               40                               45

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:07:00

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 91

Seq#:3; Line(s) 230

Seq#:5; Line(s) 373

Seq#:7; Line(s) 516

VERIFICATION SUMMARY

DATE: 04/14/2003

PATENT APPLICATION: US/09/622,419

TIME: 14:07:00

Input Set : N:\paola\US09622419.RAW.txt

Output Set: N:\CRF4\04142003\I622419.raw